

Start coding or [generate](#) with AI.

```
corpus = """
artificial intelligence is transforming modern society.
it is used in healthcare finance education and transportation.
machine learning allows systems to improve automatically with experience.
data plays a critical role in training intelligent systems.
large datasets help models learn complex patterns.
deep learning uses multi layer neural networks.
neural networks are inspired by biological neurons.
each neuron processes input and produces an output.
training a neural network requires optimization techniques.
gradient descent minimizes the loss function.
```

```

natural language processing helps computers understand human language.
text generation is a key task in nlp.
language models predict the next word or character.
recurrent neural networks handle sequential data.
lstm and gru models address long term dependency problems.
however rnn based models are slow for long sequences.
```

```

transformer models changed the field of nlp.
they rely on self attention mechanisms.
attention allows the model to focus on relevant context.
transformers process data in parallel.
this makes training faster and more efficient.
modern language models are based on transformers.
"""
```

```
import random
from collections import defaultdict
```

```
tokens = corpus.lower().replace(".", "").split()
```

```
def build_bigram(tokens):
    bigram = defaultdict(list)
    for i in range(len(tokens) - 1):
        bigram[tokens[i]].append(tokens[i + 1])
    return bigram
```

```
def build_trigram(tokens):
    trigram = defaultdict(list)
    for i in range(len(tokens) - 2):
        key = (tokens[i], tokens[i + 1])
        trigram[key].append(tokens[i + 2])
    return trigram
```

```
def generate_bigram_text(model, start_word, length=20):
    word = start_word
    output = [word]

    for _ in range(length):
        if word not in model:
            break
        word = random.choice(model[word])
        output.append(word)

    return " ".join(output)
```

```
def generate_trigram_text(model, start_words, length=20):
    w1, w2 = start_words
    output = [w1, w2]

    for _ in range(length):
        key = (w1, w2)
        if key not in model:
            break
        w3 = random.choice(model[key])
        output.append(w3)
        w1, w2 = w2, w3

    return " ".join(output)
```

```
bigram_model = build_bigram(tokens)
trigram_model = build_trigram(tokens)

print("Bigram Output:\n")
print(generate_bigram_text(bigram_model, "artificial", 25))

print("\nTrigram Output:\n")
print(generate_trigram_text(trigram_model, ("artificial", "intelligence"), 25))
```

Bigram Output:

artificial intelligence is transforming modern language models are slow for long sequences transformer models are inspired by bi

Trigram Output:

artificial intelligence is transforming modern society it is used in healthcare finance education and transportation machine lea